Scientific Research Publishing

Communications and Network, 2021, 13, 68-78
https://www.scirp.org/journal/cn
ISSN Online: 1947-3826
ISSN Print: 1949-2421

A Review on Distribution Model for Mobile Agent-Based Information Leakage Prevention

Alese Boniface Kayode[1], Alowolodu Olufunso Dayo[1], Adekunle Adewale Uthman[2]

[1]Computer Science Department, Federal University of Technolog, Akure, Nigeria
[2]Computer Science Department, Federal Polytechnic, Ede, Osun State, Nigeria
Email: adekunle.adewale@federalpolyede.edu.ng

How to cite this paper: Kayode, A.B., Dayo, A.O. and Uthman, A.A. (2021) A Review on Distribution Model for Mobile Agent-Based Information Leakage Prevention. *Communications and Network*, **13**, 68-78.
https://doi.org/10.4236/cn.2021.132006

Received: February 21, 2021
Accepted: May 24, 2021
Published: May 27, 2021

Abstract

With the continuous use of cloud and distributed computing, the threats associated with data and information technology (IT) in such an environment have also increased. Thus, data security and data leakage prevention have become important in a distributed environment. In this aspect, mobile agent-based systems are one of the latest mechanisms to identify and prevent the intrusion and leakage of the data across the network. Thus, to tackle one or more of the several challenges on Mobile Agent-Based Information Leakage Prevention, this paper aim at providing a comprehensive, detailed, and systematic study of the Distribution Model for Mobile Agent-Based Information Leakage Prevention. This paper involves the review of papers selected from the journals which are published in 2009 and 2019. The critical review is presented for the distributed mobile agent-based intrusion detection systems in terms of their design analysis, techniques, and shortcomings. Initially, eighty-five papers were identified, but a paper selection process reduced the number of papers to thirteen important reviews.

Keywords

Mobile Agent, Distribution Model, Data Leakage Detection, Data Leakage Prevention, DLP, Security, Distributed Computing

1. Introduction

The widespread use of information technology and cloud computing has increased the threats of data leakage, manipulation, and distribution. While the world is increasingly relying on cloud computing, the security and privacy concerns of the data stored and shared on the distributed network are also increasing [1]. Many enterprises had been prey to theft, loss, leakage, manipulation of sensitive business data and such kind of sensitive direct and indirect data loss of an individual or a corporate poses a great threat to the business reputation and trust of an individual [2]. Out of the reasons for data leakage and loss, the most commonly reported reason is the casual and negligent behaviour of employees while interacting with the shared data or files such as confidential information, business documents, proposals, financial statements, policies, contracts, intellectual property, and other private information [3] of a corporate which are shared through a covert channel. In such cases where the shared files are moved across a covert channel, the chances of unauthorized data access are increased as covert channels are formed through shared resources and they bypass the conventional mechanisms of data security by a legitimate user resulting in unintentional data leakage [4]. [5] reported that 90% of corporate data leakage could have been prevented through the improved strategies of data security and its leakage prevention. Deceptive and fraudulent attacks initiated through the credentials of a legitimate insider are the main contributors to the internal security breach, followed by the information leakage through the insider attack. [6] and [7] have demonstrated the importance of information security in a corporate environment, and a hybrid framework for information leakage detection and prevention is presented in [8]. Thus, it is evident that information security is a prime concerned topic in the field of IT which is aimed to provide a secure computing environment to the individual and corporate users.

A critical review is presented in this paper on the distributed model for mobile agent-based data leakage prevention. The research papers selected for this research review are from the journal and conferences which are published in 2009 and onwards. The critical research review has analyzed the distributed mobile agent-based intrusion detection and prevention systems in terms of their design, capabilities, and shortcomings. The paper also presents a review of the studies that have proposed a distribution model for mobile agents-based data leakage detection and prevention. The review includes the studies which have been proposed in 2009 and later. A discussion on these different methods is provided to analyze and compare these systems.

2. Literature Review

The handling of data is largely dependent on the category to which it belongs in term of data use, rest, or motion. As the organizations contain voluminous data which is difficult to categorize manually because of various reasons, thus it poses a great challenge for data handling and leakage. Many corporate organizations have long been trying to knob the problem of unauthorized data access, loss, and leakage as it is crucial for any organization to preserve its competitive advantage and have a consistent relationship with their customers. The common strategy of organizations is to use a Data Loss or Leakage Protection (DLP) mechanism which is a technology suite to monitor the access and transfer of confidential data and protect it from any unauthorized access. It works by detection and prevention mechanisms to stop the data breaches that occur while the data is stored on a cloud or local storage, or traveling across the network, or when it is used by the multiple users of the same network [9]. The common technological approaches for data leakage protection are depicted in **Figure 1**. It can be seen that there are four major categories of the approaches that are used for data leakage detection and prevention [10]. Designated DLP based mechanisms prevent unauthorized data access and the access and transfer of data is only possible by legitimate users. Typical methods of designated DLP solutions include machine learning-based statistical methods, pattern matching, keywords-based methods. Another common and simplest method of inspection of unauthorized access is through the use of access control mechanisms and encryption algorithms that prohibit attacks from inside or outside the environment. Some other conventional mechanisms for data leakage detection and prevention include firewall, intrusion detection procedures, and policies, and antiviruses. The standard procedures fail to work in a complex environment and are often less trustworthy in some scenarios and for this reason, other methods such as honeypot techniques, machine learning, and temporal reasoning-based mechanisms is used which includes the advanced and intelligent mechanisms to identify and detect the unauthorized attempts to attack the channel and access the data that passes through it [10]. Despite multifarious solutions, the greatest challenge is that the data loss can occur due to various reasons *i.e.* through network or storage devices, during P2P file sharing, through the internet and email transfer. Thus, data loss prevention is a multifaceted problem and DLP is not ready to use solution but requires a lot of effort to build a significant solution based on diligent preparation and constant maintenance [11].

3. Research Background

Data leakage refers to unauthorised transmission of data from within an organisation to an external destination or recipient. The types of data leaked usually include Confidential Information, Intellectual property, User Data, and Health

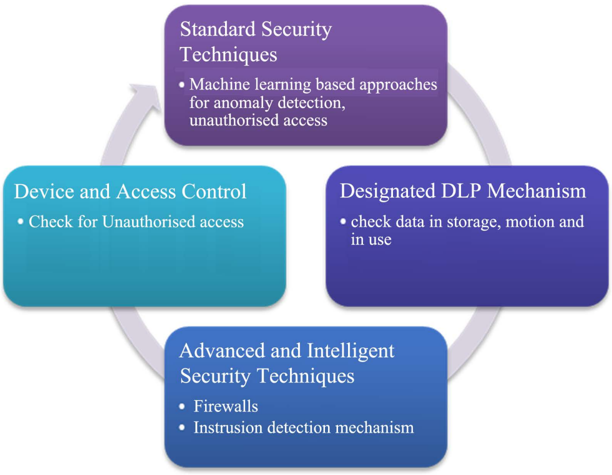
Standard Security Techniques
• Machine learning based approaches for anomaly detection, unauthorised access
Designated DLP Mechanism
• check data in storage, motion and in use
Advanced and Intelligent Security Techniques
• Firewalls
• Instrusion detection mechanism
Device and Access Control
• Check for Unauthorised access

Figure 1. Summary of the technical approaches for data leakage prevention and detection.

Records [12]. Given the present strict regulatory and legal compliance requirement on intellectual and personal data protection, organisations, including Universities, have invested a great deal of time and resources in safeguarding their information from potential unauthorised access and disclosure [13]. Many researchers have developed various counter measures to fight against data leakage issues, which are collectively known as Data Leakage Prevention (DLP) solutions. DLP can be defined as any systems or tools that identify, monitor, and protect data [14]. Examples of such data include:

1) Data in Motion—Data transmitted on wire or wirelessly.

2) Data in Use—Data at the endpoints of the network such as information stored on portable hard disks.

3) Data at Rest—Data that resides in files system, databases and other storage methods.

Some of the commonly available DLP Systems are:

1) Network DLP: Network DLP is designed to detect any leakage incidents related to data in motion, by detecting if particular important data files are being transferred through networks. This kind of DLP devices usually attached to equipment such as routers, switches supports multiple protocols such as HTTP, FTP, P2P and SMTP.

2) Endpoint DLP: Endpoint DLP products are agents or software that usually reside on end user terminals such as mobile devices and laptops. The common use of Endpoint DLP is to prevent users from storing sensitive information on removable media devices such as USB flash drives and CD/ROM discs and to protect against unauthorised transmission of sensitive information when a user is not connected to public free Wi-Fi spot. An Endpoint DLP software can also utilise disk encryption, which prevents unauthorsied access to information on a lost or stolen laptop.

3) Embedded DLP: Embedded DLP is planted within specific applications to effectively monitor the data outflows, identify keywords or related patterns belong to sensitive information and block any suspicious data leakage attempts. For instances, scanning and rejecting outgoing e-mails for sensitive keywords or attachments, restricting printing of copyrighted softcopy documents.

Some of the key benefits of DLPs are Prevent Data Leakage, Reduce Cost of Investigation and Damage to Reputation, Facilitate Early Risk Detection and Mitigation, and Increase Comfort Level of Senior Management [15].

4. Review of Existing Work

DLP is a multifaceted problem that deals with what is responsible for the data loss and what should be done to avoid this situation. Many methods for data leakage detection and prevention have been presented in the past which in one way or the other provides the solution of the problem or lack in some way to be an absolute solution. Brief literature of these solutions for data leakage detection and prevention is described in the following subsections.

4.1. Data Leakage Detection

Digital watermarks have long been used for authentication and to ensure the integrity of the carrier signal. Moreover, they are also being used to identify the identity of the owner, thus, they identify the source of the data forgery or leakage and prevent the illegal manipulation of data. For this reason, watermarking approaches are used to detect data manipulation and leakage. But the problem with the watermarking techniques is that when the watermark is embedded in the source, the source does not remain in its original form and is modified. On the other hand, if the source cannot be modified, a watermark cannot be added in it [16]. SELinux and Colored Linux are the two watermarking based methods used for data leakage detection. Colored Linux is a successor and is formed by modifications in SELinux. Colored Linux generates blind watermarks based on the tags granted to a file (permission of file access) by a filesystem. Thus, when a file is accessed, it examines the tag with the watermark and provides access accordingly. Thus, unauthorized accesses are expelled if the tag does not match the watermark of the file else SELinux provides the access control to the user accessing the file. Such mechanisms work well for close systems where one system is not connected to the external machines and the modified operating system executes on every system. However, it does not work well for an open system where a system is connected to multiple systems with non-colored (modified) operating systems. In such cases, insider attacks can occur through the covert channels leaving the files modified and un-noticed [17]. A novel hybrid method utilizing the mobile agents and the colored watermarking scheme is proposed in [17] which defines a strong relationship between the security tag and a blind watermark and effectively tackles the insider and covert channel attacks.

Software agents are the individual autonomous programs that work on behalf of other entities to achieve some goals. [17] proposed a software agent-based Information Leakage Detection (ILD) which consists of seven types of software agents and whose inter-agent communication is governed by FIPA Agent Communication Language (ACL). The seven types of agents include control, permission, environment, watermarking, monitor, queue, and detection agents each with their own set of jobs. The proposed agent-based mechanism of [17] requires less administrative effort (only for installation and configuration in the start) and each mobile agent has a certain set of abilities that help in the identification of data leakage.

In another study, the identification of the source of the data leakage is reported to be done by the data provenance technology which identifies the legitimate owners of the data (found in the database) and thus, the sources leaking the data are identified [18]. Another technique has been proposed that make use of fake data insertion to improve the probability of identification of data leakage [19]. A slight modification in the technique of [19] is the addition of the user's guilt probability which improves the file distribution algorithm. One of the prime reasons for the data leakage is because of human involvement *i.e.* the intent of the malice workers. It is reported that 11% of the employees sold the business information for profit through unauthorized access [20]. For this reason, two models have been presented in [20] *i.e.* watcher model and guilt model. Watcher model analyses the unauthorized access of a corporate employee and the guilt model utilizes the guilt probability to analyze whether the suspected user leaks this file to some third party. These two models are proposed to calculate the likelihood of user involvement in the data leakage by observing the intersection between the leaked data and the data access of the suspected user. Vasileios *et al.* [21] developed iLeak system based on the components of the commodity operating system to detect the sensitive data loss for the personal system.

4.2. Data Leakage Prevention

While there are dozens of algorithms for the identification of data leakage, it is the priority to choose methods that utmost mitigates the situations leading to the problem of data leakage. The criteria for the relevance of data attacks and the characterization of the data loss incidents are defined in [22] which builds a notion for data and information leakage prevention (DLP & ILP) in information security. To protect the organizational information against the inter and intra information disclosure attacks, a trust management model is presented [23]. [24] presents the trustworthiness model for the data leakage prevention in which the user's trustworthiness is taken into account which is calculated from the user's history, then his intentional leakage behavior is calculated from his trustworthiness and file access history. The overlap between the two indicates the user's intentional leak behavior. Also, the distribution evaluates the vulnerability of the user’s platform. Based on this collective risk evaluation, the decision is taken by the distributor to whether or not to distribute the files to these users. [25] proposed a model in Hadoop based master/slave architecture for data leakage detection and prevention. The model utilizes Reliability Checker (RC), the least reliable agent (LRA), and Data Leakage Avoider (DLA) algorithms. The least reliable agent is calculated by the LRA algorithm and its output is sent to DLA to prevent the allocation of data to these nodes.

Figure 2 represents the hybrid framework that combines the mobile agents and DLP for the detection and prevention of data leaks by designing this framework on the kernel-level filesystem of the operating system. It is designed to overcome the weaknesses which arrive while monitoring the activities of data leakage. This proposed mechanism is based on the kernel level file subsystem which is previously available in the commodity operating system to identify the potential leaks of the sensitive data. This hybrid framework reduces the administrative work and provides the ability to modify and add data loss detection capabilities and modularize them employing a control agent. This approach is reported to give a lesser false-positive rate because of its ability to identify the unknown attacks. **Table 1** summarizes the solutions proposed for data leakage detection and prevention along with their findings.

Table 1. Review of the existing distribution models for data leakage prevention.

Year & Reference	Method	Findings	Limitations
2009 [13]	Proposed a mobile agent-based mechanism combined with the coloring *i.e.* robust watermarking to identify the information leakage sources.	The proposed method effectively identifies the potential leakage sources from both the covert channel and the insiders.	Implemented primarily through modification of the SELinux kernel modules. Experimental details and results of the host-resident agents in technique were not represented.
2010 [20]	Proposed a model to assess the malicious and honest users.	The model effectively classifies the malicious and honest users and prevent the distribution of files to them thus, preventing the data leakage.	The model uses a single classification technique to classify malicious and honest uers.
2010 [17]	Proposed a system named as iLeak for personal data loss detection and is lightweight as compared to other proposed systems.	This lightweight system effectively prevents the inadvertent data leaks and produces overhead of 4% for the protected systems and applications.	Detection approach relies on keywords for representing sensitive information, there is a chance for false alerts.
2011 [23]	Proposed an algorithm for automatic classification of corporate documents as sensitive or not-sensitive.	Effectively classifies the sensitive corporate sensitive documents and works well on big data.	Most of the works studied employed the used of a single machine learning technique (SVM) for document classifiers.
2012 [16]	Developed two models *i.e.* watcher and guilt model. Watcher model identifies the unauthorized access and guilt model defines the probability of identifying the guilty distribution parties.	Assesses the probability of an agent to be responsible for the data leakage.	The models developed were to evaluated.
2013 [12]	Makes use of the user's guilt probability to define a file allocation plan.	Effectively identifies the leak source and provides a file allocation plan.	Provide little or no support for alert handling.
2015 [18]	Defines the criteria for characterizing the significance and relevance of data attacks and advanced criteria for characterizing the data loss incidents.	Complete protection against the data loss in a corporate sector is impossible as human involvement is a key decisive factor in the data-information leakage prevention.	No practical/functional Information Leakage Detection and Prevention system had been implemented for a distributed system.
2015 [24]	Proposed a dynamic three-phase data leakage detection scheme.	The proposed method efficiently identifies the anomalous behavior, detects and classifies the data leakage resources.	The result presented by the author indicated that C4.5 is the best machine learning techniques but C4.5 does not work very well on a small training set.

Continued

2016 [21]	Hadoop based master/slave architecture that utilizes Least Reliable Agent (LRA) algorithm for data leakage prevention and Reliability Checker (RC) and Data Leakage Avoider (DLA) algorithms to prevent allocation of data to the less reliable and suspected leakage nodes.	Effectively reduces the data leakage.	The purpose of this study is descriptive.
2016 [8]	Proposed a hybrid framework for data leakage detection and Prevention.	Effectively identifies the insider attacks and anomalous behavior.	This research work fails to show in detail the anomaly-based techniques adopted.
2017 [9]	Proposed BROSMAP: A Novel Broadcast Based Secure Mobile Agent Protocol for Distributed Service Applications.	The proposed system provides protection from man in the middle, replay, repudiation, and modification attacks.	The application does not incorporate a trust model to help users evaluate the honesty and behavior of service providers.
2018 [10]	Proposed a data leakage prevention method based on the reduction of confidential and context terms for smart mobile devices.	It presents a pruning method based on the attribute reduction method of rough set theory.	Does not provide sufficient solution against intentional leakages.
2019 [11]	Proposed an agent based information security framework for hybrid cloud computing.	The results confirm that proposed framework could be used for information security in cloud computing environment.	No proper evaluation was carried out to check for the validity and reliability of the proposed framework.

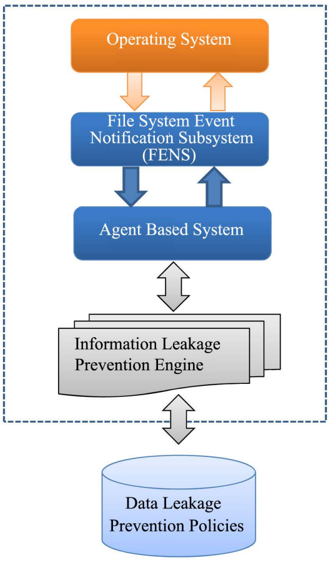
Operating System
File System Event Notification Subsystem (FENS)
Agent Based System
Information Leakage Prevention Engine
Data Leakage Prevention Policies

Figure 2. The architecture of mobile agent-based information leakage detection and prevention.

4.3. Limitations in the Existing Distribution Models of Mobile Agent-Based Data Leakage Detection and Prevention

While a dozen solutions have been proposed to detect and prevent data leakage, no solution has found to provide complete defense against the problem [26]. Therefore, some limitations are found in the existing solutions. One of the shortcomings that have been identified is the lack of a generic framework for data leakage detection and prevention and therefore, many solutions have been proposed but none is found to have a generic workflow for the basic understanding of the problem. Another limitation is the lack of a practical and functional framework for a distributed system for the problem of data leakage detection and prevention. Moreover, the existing proposed solutions have done simulations on the Linux operating system whereas a large population of IT is of people who use the Windows operating system. Also, there is limited use of machine learning techniques has been seen *i.e.*, only Support Vector Machine (SVM). Lastly, it is also observed that in some cases no support or sustenance is provided for the scenarios of data leakage alerts.

In most of the papers reviewed, it showed that

- No practical/functional Information Leakage Detection and Prevention system had been implemented for a distributed system.
- The development of a generic framework for data leakage detection and prevention is lacking.
- Most of the works studied employed the use of a single machine learning technique (SVM) for document classifiers.
- Most of the works done in DLP are simulated on Linux operating systems (instead of Windows OS with the largest number of users).
- A drawback of existing literatures is that they provide little or no support for alert handling.

This research is therefore motivated by the need to overcome these limitation via a novel integration of Mobile-Agent system with Information Leakage Detection and Prevention system build on operating system kernel-level file system; as a way of using their respective unique strength for enhancement and elimination of weaknesses in the task of monitoring system activities against information leaks.

5. Conclusion

The technological advancement and prevailing use of cloud and distributed computing have increased the security issues related to the data and IT and the threats associated with the data leakage in such an environment have also increased manifolds. This research review has critically analyzed the solutions that have been proposed in the last decade for data leakage detection and prevention. The purpose of various studies was descriptive and many studies did not provide a significant and effective solution for data leakage detection and prevention for the distributed systems. While various proposed solutions suggested and improved one or the other aspect to prevent data leakage and identify the malicious sources and users but none of the solutions is complete and absolute. However, some solutions that have been proposed lack in theory though they have performed well in some scenarios. Therefore, it is concluded that DLP systems require significant research to be developed for real-world systems. Furthermore, it is not a plug-and-play solution to prevent data leakage and requires continuous research and maintenance until an acceptable and comprehensive framework is developed.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] British Broadcasting Corporation (BBC) (2014) Target Data Theft Affected 70 Million Customers.

[2] Verizon (2015) Data Breach Investigations.

[3] Alneyadi, S., Sithirasenan, E. and Muthukkumarasamy, V. (2016) A Survey on Data Leakage Prevention Systems. *Journal of Network and Computer Applications*, **62**, 137-152. https://doi.org/10.1016/j.jnca.2016.01.008

[4] Alneyadi, S., Sithirasenan, E. and Muthukkumarasamy, V. (2015) Detecting Data Semantic: A Data Leakage Prevention Approach. 2015 *IEEE Trustcom/BigDataSE/ISPA*, Helsinki, 20-22 August 2015, 910-917. https://doi.org/10.1109/Trustcom.2015.464

[5] Margathavall, R.M.P., Pramila, V., Priya, R. and Abiram, P. (2016) Preserving Sensitive Data By Data Leakage Prevention Using Attribute-Based Encryption Algorithm. *International Journal of Emerging Technology in Computer Science & Electronics*, **21**, 3.

[6] Greenwald, G., MacAskill, E. and Poitras, L. (2013) Edward Snowden: The Whistleblower behind the NSA Surveillance Revelations. *The Guardian*, **9**, 2.

[7] Mortensen, M. (2014) Who Is Surveilling Whom? Negotiations of Surveillance and Sousveillance in Relation to Wikileaks' Release of the Gun Camera Tape *Collateral Murder*. *Photographies*, **7**, 23-37. https://doi.org/10.1080/17540763.2014.896144

[8] Costante, E., Fauri, D., Etalle, S., Den Hartog, J. and Zannone, N. (2016) A Hybrid Framework for Data Loss Prevention and Detection. 2016 *IEEE Security and Privacy Workshops*, San Jose, 22-26 May 2016, 324-333. https://doi.org/10.1109/SPW.2016.24

[9] Ren, L. (2013) DLP Systems: Models, Architecture and Algorithms. 19-21.

[10] Shabtai, A., Elovici, Y. and Rokach, L. (2012) A Survey of Data Leakage Detection and Prevention Solutions. Springer Science & Business Media, Boston. https://doi.org/10.1007/978-1-4614-2053-8

[11] Raman, P., Kayacık, H.G. and Somayaji, A. (2011) Understanding Data Leak Prevention. 6*th Annual Symposium on Information Assurance*, Albany, 7-8 June 2011, 27.

[12] Yin, F., Yu, R., Wang, L. and Ma, X. (2013) A Distribution Model for Data Leakage Prevention. *Proceedings of* 2013 *International Conference on Mechatronic Sciences*,

Electric Engineering and Computer, Shenyang, 20-22 December 2013, 2617-2620. https://doi.org/10.1109/MEC.2013.6885474

[13] Lee, Y.-C., Bishop, S., Okhravi, H. and Rahimi, S. (2009) Information Leakage Detection in Distributed Systems Using Software Agents. 2009 *IEEE Symposium on Intelligent Agents*, Nashville, 30 March-2 April 2009, 128-135. https://doi.org/10.1109/IA.2009.4927510

[14] Buneman, P. and Tan, W.-C. (2007) Provenance in Databases. *Proceedings of the* 2007 *ACM SIGMOD International Conference on Management of Data*, Beijing, June 2007, 1171-1173. https://doi.org/10.1145/1247480.1247646

[15] Papadimitriou, P. and Garcia-Molina, H. (2009) A Model for Data Leakage Detection. 2009 *IEEE* 25*th International Conference on Data Engineering*, Shanghai, 29 March-2 April 2009, 1307-1310. https://doi.org/10.1109/ICDE.2009.227

[16] Ajay Kumar, J. and Rajani Devi, K. (2012) An Efficient and Robust Model for Data. *Leakage Detection System*, **3**, 91-95.

[17] Kemerlis, V.P., Pappas, V., Portokalidis, G. and Keromytis, A.D. (2010) iLeak: A Lightweight System for Detecting Inadvertent Information Leaks. 2010 *European Conference on Computer Network Defense*, Berlin, 28-29 October 2010, 21-28. https://doi.org/10.1109/EC2ND.2010.13

[18] Hauer, B. (2015) Data and Information Leakage Prevention within the Scope of Information Security. *IEEE Access*, **3**, 2554-2565. https://doi.org/10.1109/ACCESS.2015.2506185

[19] Srivatsa, M., Balfe, S., Paterson, K.G. and Rohatgi, P. (2008) Trust Management for Secure Information Flows. *Proceedings of the* 15*th ACM Conference on Computer and Communications Security*, Alexandria, October 2008, 175-188. https://doi.org/10.1145/1455770.1455794

[20] Fan, Y., Wang, Y., Wang, L. and Yu, R. (2010) A Trustworthiness-Based Distribution Model for Data Leakage Prevention. *Wuhan University Journal of Natural sciences*, **15**, 205-209. https://doi.org/10.1007/s11859-010-0305-7

[21] Praveen, D.Y.S.S., Suba Rao, D. and Kumar, A. (2016) A Novel Model for Data Leakage Detection and Prevention in Distributed Environment. *International Journal of Engineering and Technical Research*, **4**, 9.

[22] Revathi, Y. ad Kumar, D.S.M. (2016) Review on Importance and Advancement in Detecting Sensitive Data Leakage in Public Network. *International Journal of Engineering Research and General Science*, **4**, 263-265.

[23] Hart, M., Manadhata, P. and Johnson, R. (2011) Text Classification for Data Loss Prevention. *International Symposium on Privacy Enhancing Technologies Symposium*, Waterloo, 27-29 July 2011, 18-37. https://doi.org/10.1007/978-3-642-22263-4_2

[24] Maldonado, C. (2015) Data Leakage Detection Using dynamic Data Structure and Classification Technique. *INGE CUC*, **11**, 79-84.

[25] Faiz, M.F., Arshad, J., Alazab, M. and Shalaginov, A. (2020) Predicting Likelihood of Legitimate Data Loss in Email DLP. *Future Generation Computer Systems*, **110**, 744-757. https://doi.org/10.1016/j.future.2019.11.004

[26] Graef, I., Husovec, M. and Purtova, N. (2018) Data Portability and Data Control: Lessons for an Emerging Concept in EU Law. *German Law Journal*, **19**, 1359-1398. https://doi.org/10.1017/S2071832200023075